Insa1

TIMING ATTACK RESISTANT CRYPTOGRAPHIC SYSTEM

The present invention relates to the field of cryptographic systems and in particular to a method and apparatus for resisting timing attacks on a cryptographic system.

BACKGROUND OF THE INVENTION

Cryptographic systems generally owe their security to the fact that a particular piece of information is kept secret without which it is almost impossible to break the scheme. This secret information must generally be stored within a secure boundary, making it difficult for an attacker to get at it directly. However, various schemes or attacks have been attempted in order to obtain this secret information. One of these is the timing attack.

By way of background current public key cryptographic schemes such as RSA and elliptic curve (EC) operate over mathematical groups F_p^* and $E(Fq)$ respectively. The group operations, called multiplication modulo p, in RSA, and addition of points in EC are repeated in a particular way to perform a scalar operation. In RSA the operand is called an exponent, the operation is called exponentiation and the method of multiplying is commonly known as repeated square-and-multiply. Thus given a number $a \in F_p$ and an integer $0 \leq k < p$, the exponent, whose binary representation is $k = \Sigma_{i=o}^{t} k_i 2^i$ a value $a^k \bmod p$ may be calculated by repeated use of the square-and-multiply algorithm. Similarly given $g(x) \in F_p m$ and an integer $0 \leq k \leq p^m - 1$ then $g(x)^k \bmod f(x)$ may be calculated by this method.

On the other hand, in EC the operand is a scalar multiplier, the operation is called scalar multiplication of a point, and the method is known as double-and-add. Thus if a is a positive integer and P is an elliptic curve point then aP may be obtained by the double-and-add method. Both these methods are well known in the art and will not be discussed further.

In RSA, half of all exponentiation operations use a private key. Whereas in EC all scalar multiplications use either a long term private key or a session private key. In each of these cases, the private key is safe due to the difficulty of reversing the exponentiation or multiplication operation as the case may be. This is based on the discrete log problem

or the difficulty of integer factorization. As mentioned earlier, an attacker once in possession of the private key (either long term or session) is able to forge signatures and decrypt secret messages for the attacked entity. Thus it is paramount to maintain the secrecy or integrity of the private key in the system.

Many techniques have been suggested to obtain the private key. The encryption operations are performed either in a special purpose or general purpose processor operating in a cyclic manner. Recent attack methods proposed in open literature have been based on timing analysis of these processors or in other words timing analysis of 'black box' operations. In one instance an attacker by capturing the instantaneous power usage of a processor throughout a private key operation obtains a power signature. The power signature relates to the number of gates operating at each clock cycle. Each fundamental operation as described in the preceding paragraph generates a distinct timing pattern. Other methods exist for obtaining a power signature than instantaneous power usage.

Laborious but careful analysis of an end-to-end waveform can decompose the order of add-and-double or square-and-multiply operations. Either a double or square must occur for each bit of either the exponent or scalar multiplier respectively. Therefore, the places where double waveforms are adjacent each other represent bit positions with zeros and places where there are add patterns indicate bits with ones. Thus these timing measurements can be analyzed to find the entire secret key and thus compromise the system. Thus there is a need for a system which minimizes the risk of a successful timing attack.

SUMMARY OF THE INVENTION

This invention thus seeks to provide a cryptographic system where cryptographic operations are performed by a processor in a constant period of time irrespective of the operation being performed whereby a constant amount of time is required for the processing of each bit scalar or a exponent regardless of its value.

A method for determining a result of a group operation performed an integral number of times on a selected element of the group, said method comprising the steps of:

(a) representing said integral number as a binary vector;

(b) initializing an intermediate element to the group identity element;

(c) selecting successive bits, beginning with a left most bit, of said vector and for each of said selected bits;

(i) performing said group operation on said intermediate element to derive a new intermediate element;

(ii) replacing said intermediate element with said new intermediate element;

(iii) performing said group operation on said intermediate element and an element, selected from the group consisting of:

said group element if said selected bit is a one; and

an inverse element of said group element if said selected bit is a zero;

(iv) replacing said intermediate element with said new intermediate element;

(d) performing said group operation on said intermediate value and said inverse element if said last selected bit is a zero; and replacing said intermediate element therewith, to obtain said result, whereby each of the bits of said integral is processed with substantially equal operations thereby minimizing timing attacks on said cryptographic system.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features of the invention will now be described by way of example only with reference the accompanying drawings in which:

Figure 1 is a schematic representation of a data communication system;

Figure 2 is a schematic representation of a binary number recoding scheme;

Figure 3 is a state machine representation for compiling a scalar multiple of a point;

Figure 4 is a pseudocode implementation of the state machine of Figure 3;

Figure 5 is a non-state machine pseudocode implementation;

Figures 6 and 7 are respectively a pseudocode and state-machine implementation for a square and multiply scheme; and

Figure 8 is a generalized schematic diagram of an apparatus for implementing the method according to an embodiment of the invention.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Referring therefore to Figure 1, a secure data communication system 10 includes a pair of correspondents, designated as a sender A(12), and a recipient B(14), who are connected by a communication channel 16. Each of the correspondents A and B (12,14) includes an encryption unit 18,20 respectively that may process digital information and prepare it for transmission across the channel 16.

Generally, the sender A assembles a data string, which includes amongst others the public key y of the sender, a message m, the sender's short-term public key k and signature S of the sender A. When assembled the data string may be forwarded to the intended recipient B, who then verifies the signature using A's public key. This public key information may be obtained from a certification authority (CA) 24 or sometimes is sent with the message.

For example, in RSA public key encryption, B encrypts a message m for A, which A decrypts. To encrypt the message m, B obtains A's public key (n, e) and represents the message as an integer in the interval [0, m-1]. Next B computes $c = m^e \bmod n$ and sends the cipher text c to A. The entity A then recovers the message m from c by using its private key d to recover $m = c^d \bmod n$. The calculation $c^d \bmod n$ (modulo exponentiation) may be performed by using a well known square and multiply algorithm. Similar computations are required for signing in RSA, and for signing, encryption and decryption in discrete log systems such as ECC.

The values used in the computations are expressed as bit vectors, which are then manipulated by the encryption processor in accordance with a particular encryption scheme being used. Thus, referring to figure 2, any bit vector ending in a one can be recoded into an all non-zero vector of +1 and -1 terms. For elements ending in a zero, it is necessary to either maintain the final zero in the recoded form or to perform a further -1 action as a correction after the final zero is processed by the previous rules. As will be apparent later, it is preferable to do the corrective action.

Referring to figure 3 a state machine implementation of an algorithm for computing an integer multiple b of a point P on an elliptic curve is shown generally by numeral 30. In this embodiment, a bit vector b to be processed is fed into the state machine 30. The result of the processing is the value bP where P is a point having

coordinates (x,y) on an elliptic curve. The state machine 30 is initialized with a counter *i* set to the number of bits N in the vector b, and an intermediate value Q stored in a register. Entry into state H0 of the state machine occurs when the first bit (MSB) of b is encountered (not the first non-zero bit). In state H0 the contents of the intermediate Q is doubled and the base point P is added thereto. The bits are sequentially processed causing transitions either back to the current state H0 or to a next state H1. In state H1, Q is doubled and the base point P is subtracted. When the next bit is a zero, the next state is always H1. When the next bit is a one, the next state is always H0. Whenever the current state is H0, an add (+P) occurs, otherwise a subtract (-P) occurs. Once all bits are processed, control exits to the DONE state with the result bP. If the exit condition is encountered while in H0, it is necessary to subtract (-P) a final time from the result.

Implementation of the state machine in pseudocode form requires the availability of a general purpose storage bit for the H state. Figure 4 shows a pseudocode implementation of the state machine where general purpose data and control mechanisms are available. Timing attacks are prevented only if the execution time and power are identical for all possible execution paths through the loop. In this implementation, this requires that the time for the execution path through lines L5, L6 and L7 be the same as that through lines L5, L8 and L9. This implies that the branch instruction (IF) must execute in the same time for false and true conditionals.

Furthermore, the add and subtract operations must take the same time. By adding in the negative value of P, this is more obviously possible. Note that the execution time of all other lines in the algorithm are non-critical to the timing attack resistance, with the exception of lines L14 and L15. These two lines, if executed will increase the total time required, thus revealing that the final H state was zero. This is equivalent to revealing that the final bit of the scalar was a zero. This is unavoidable, but only reveals a single bit of the scalar multiple.

On application specific computing devices, it is likely that there are no general purpose data storage areas nor general purpose assignment and test operators. In this case, the H state control cannot be added in a usual way. Instead, it is necessary to encode the state by branching through distinct code paths.

Referring to Figure 5 shows a 'state-less' pseudocode implementation. Each code execution path corresponds to a distinct state. While this will in general cause code

expansion, here there are only two short paths required. Here, as above, timing attacks are prevented only if the execution time and power are identical for all possible execution paths through the loop. In this implementation, this requires that the time for the execution path through lines LL3, LL4 and LL5 be the same as that through lines LL6, LL7 and LL8. In addition, path LL9, LL10 and LL16 must execute in the same time and power as path LL9 and LL11. This will be true in architectures where conditional branch instructions take the same time to branch (to a new location) as to fall through (to the following location). Note that lines LL11, LL12 and LL13 are equal in execution time and power to lines LL16, LL17 and LL18. This is also necessary to prevent timing attacks. Otherwise, H state information would be revealed. As in the previous implementation, a final corrective subtract is required line LL14 when the loop terminates through the H0 path. This again reveals the final H path, or equivalently the final bit of the scalar.

Referring to Figure 6 a pseudocode implementation of a square and multiply operation is shown by numeral 60, while a corresponding state machine implementation is shown in Figure 7. In this implementation, the exponent bit vector b is fed into the state machine 60. The result of the processing is to compute a value M^b. Once again a counter is initialized to the length N of vector b and an accumulator Q is initialized to one. Entry into state H0 of the state machine begins with the MSB of bit vector b. As previously, the bits are sequentially processed causing transitions either back to the current state H0 or to a next state H1. In state H0 the contents of the accumulator Q is squared and multiplied by the base M. In state H1 the accumulator is also squared and then divided by the base M. If the next bit is a zero, the next state is always H1, whereas if the next bit is a one the next state is always H0. Whenever the current state is H0, a multiply occurs otherwise a divide occurs. As previously, once all bits are consumed, controls exits to the DONE state. If the exit condition is encountered while in H0, it is necessary to divide a final line by the base M.

Figure 8 shows a generalized processor implementation in which a state controller 82 is programmed to run the estate code as described earlier. A modulo arithmetic and/or finite field computation unit 84 is provided for computing the square and multiplication and the point additions/subtractions respectively. A counter 86 and a register b is coupled to the state controller for timing the sequential operation of the controller 82.

While the invention has been described in connection with a specific embodiment thereof and in a specific use, various modifications thereof will occur to those skilled in the art without departing from the spirit of the invention as set forth in the appended claims.

The terms and expressions which have been employed in the specification are used as terms of description and not of limitations, there is no intention in the use of such terms and expressions to exclude any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the claims to the invention.